Generalizations of Ho-Lee's binomial interest rate model I: from one- to multi-factor

Jirô Akahori[1], Hiroki Aoki[2], and Yoshihiko Nagata[3]

[1]Department of Mathematical Sciences,
Ritsumeikan University
1-1-1 Nojihigashi, Kusatsu, Shiga 525-8577, Japan
e-mail: akahori@se.ritsumei.ac.jp

[2]Department of Mathematics,
Tokyo University of Science
Noda, Chiba, 278-8510, Japan
e-mail: aoki_hiroki@ma.noda.tus.ac.jp

[3]Risk Management Department,
Mizuho Trust & Banking Co. Ltd.
1-2-1 Yaesu, Tokyo, 103-8670, Japan
e-mail: yoshihiko_nagata@tz.mizuho-tb.co.jp

This version: June 2006

Abstract

In this paper a multi-factor generalization of Ho-Lee model is proposed. In sharp contrast to the classical Ho-Lee, this generalization allows for those movements other than parallel shifts, while it still is described by a recombining tree, and is stationary to be compatible with principal component analysis. Based on the model, generalizations of duration-based hedging are proposed. A continuous-time limit of the model is also discussed.

Contents

Keywords: Ho-Lee model, duration, multi-factor, recombining tree, stationarity, forward rate, drift condition
Mathematical Subject Classification 2000 :91B28 60G50
Journal of Economic Literature Classification System:G12
This research was supported by Open Research Center Project for Private Universities: matching fund subsidy from MEXT, 2004-2008 and also by Grants-in-Aids for Scientific Research (No. 18540146) from the Japan Society for Promotion of Sciences.

1 Introduction

1.1 The aim of the present paper/missing rings between Ho-Lee and HJM

It is almost thirty years since Vasicek's paper on term structure of interest rates [25] was published, and is almost twenty years since the paper on binary interest rate by Ho and Lee [13] appeared. Most of the papers on interest rate modeling in these 20-30 years were in anyway coming from either Vasicek or Ho-Lee.

Vasicek initiated so called *spot rate models*, followed by Cox, Ingersoll and Ross [8] and Hull and White [16] among others, and its multi-factor generalizations were proposed, for example, by Longstaff and Schwartz [21], and by Duffie and Kan [10] they are unified as Affine Term Structure Models (ATSM). Recently modifications of ATSMs like Quadratic Term Structure Models (QTSM) are extensively studied (see, e.g., Chen, Filipović, and Poor [7] and the references therein). In a word, they are *alive*.

After Ho-Lee, on the other hand, there was a major progress made by Heath, Jarrow and Morton [11] and [12]; the shift to forward rate models in a continuous-time framework[1]. The success of HJM was so great that Ho-Lee model almost lost its position as a teacher of HJM. In many textbooks Ho-Lee model is treated as an almost trivial example of HJM or even worse, spot rate models.

Starting from a careful study of the original Ho-Lee, the paper will shed lights on what are missed and lost in the way between Ho-Lee and HJM or after HJM. We will liberate Ho-Lee model from the rule of parallel shifts by shifting to forward rate modeling. In contrast to HJM, we preserve the traditions of recombining tree and stationarity, and hence consistency with principal component analysis.

As for applications, we will concentrate on generalizations of the duration-based hedging.

Most of our results presented here could be already well-known among practitioners, but in anyway they have not been explicitly stated. We do this to recover missing rings between practitioners and academics. Though the style of the presentation is that of mathematicians, we have carefully avoided

[1] Ho and Lee themselves proposed a multi-factor generalization within a recombining-tree framework [14], but actually it is not that much so far.

too abstract mathematics.

1.2 Sensitivity analysis of interest rate risks /What do we mean by "multi-factor"?

> The key assumption underlying the duration-based hedging scheme is that all interest rate change by the same amount. This means that only parallel shifts in the term structure are allowed for. (J. Hull [15])

The present value of a stream of cash flow is a function of the current term structure of interest rate. The *duration* and the *convexity* are the most classical sensitivity criteria to handle the interest rate risks (see e.g. [15, section 4.8 & 4.9]).

Let us review the context. Suppose that, for $T = T_1, \ldots, T_k$ in the future, we will have $\mathbf{CF}(T)$ of deterministic cash flow, while the term structure of interest rates are given as $\mathbf{r}(T)$. Then the present value at time t $(< T_1)$ of the cash flow is

$$\mathrm{PV}_t \equiv \mathrm{PV}(\mathbf{r}(T_1), \cdots, \mathbf{r}(T_k); t) = \sum_{l=1}^{k} \mathbf{CF}(T_l) e^{-\mathbf{r}(T_l)(T_l - t)}. \tag{1.1}$$

The *duration* in the usual sense

$$(\text{duration}) \equiv (\text{dur}) := \frac{1}{\mathrm{PV}} \sum_{l=1}^{k} (T_l - t) \mathbf{CF}(T_l) e^{-\mathbf{r}(T_l)(T_l - t)}$$

can be obtained as a derivative of the multivariate function $\log \mathrm{PV}$ in the direction of $\mathbf{ps} \equiv (1, \ldots, 1)$:

$$D_{\mathbf{ps}} := -\partial_{\mathbf{ps}} (\log \mathrm{PV}) = -\frac{1}{\mathrm{PV}} \lim_{\epsilon \to 0} \epsilon^{-1} \{ \mathrm{PV}(\mathbf{r} + \epsilon\, \mathbf{ps}; t) - \mathrm{PV}(\mathbf{r}; t) \} = (\text{dur}). \tag{1.2}$$

The *convexity* also comes from the second order derivative

$$D_{\mathbf{ps}}^2 := \frac{1}{\mathrm{PV}} \partial_{\mathbf{ps}}^2 (\mathrm{PV}) = -\frac{1}{\mathrm{PV}} \sum_{l=1}^{k} (T_l - t)^2 \mathbf{CF}(T_l) e^{-\mathbf{r}(T_l)(T_l - t)}. \tag{1.3}$$

Suppose that the random dynamics of the interest rates over the period $[t, t+\Delta t]$ is described by a real valued random variable Δw_t:

$$\Delta \mathbf{r}_t := \mathbf{r}_{t+\Delta t} - \mathbf{r}_t = \mathbf{ps}\,\Delta w_t + \mathbf{trd}\,\Delta t, \tag{1.4}$$

where $\mathbf{trd}$ is a k-dimensional vector which describes the deterministic "trend". Then we have the following (Itô type) Taylor approximation that describes the random evolution of the present value:

$$\begin{aligned}\frac{\Delta(\mathrm{PV})}{\mathrm{PV}} &:= \frac{\mathrm{PV}(\mathbf{r}+\Delta\mathbf{r}_t; t+\Delta t) - \mathrm{PV}(\mathbf{r}; t)}{\mathrm{PV}} \\ &\simeq -D_{\mathbf{ps}}\,\Delta w_t + \frac{1}{2} D^2_{\mathbf{ps}}\,(\Delta w_t)^2 \\ &\qquad + \left\{ \frac{\partial_t(\mathrm{PV})}{\mathrm{PV}} - D_{\mathbf{trd}} \right\} \Delta t,\end{aligned} \tag{1.5}$$

where $D_{\mathbf{trd}}$ is defined in the same way as (1.2).

From a viewpoint of risk management, the expression (1.5) implies that, if we could choose $\mathbf{CF}$ so as to have $D_{\mathbf{ps}} = 0$ (and $D^2_{\mathbf{ps}} = 0$), then for the short time period the portfolio is riskless. Roughly, the duration corresponds to so called *delta* and the convexity, *gamma*.

These criteria, however, work well only if the movements of the term structure of interest rate are limited to parallel shifts, as pointed out in the last part of Chapter 4 of [15], which is quoted above. To have richer descriptions, we want to make the dynamics **multi-factor**: typically we want to assume

$$\Delta \mathbf{r}_t = \sum_{j=1}^{n} \mathbf{x}^j\,\Delta w^j_t + \mathbf{trd}\,\Delta t, \tag{1.6}$$

where $\mathbf{x}^j j$'s are k-dimensional vectors and Δw^j_t's are real valued random variables with

$$\mathbf{E}[\Delta w^j_t] = 0 \text{ for all } j, \text{ and } \mathrm{Cov}(\Delta w^i_t, \Delta w^j_t) = \begin{cases} \Delta t & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{1.7}$$

Here n is the number of *factors*. If we write $\Delta \mathbf{w} := (\Delta w^1, ..., \Delta w^n)$, the equation (1.6) is rewritten as

$$\Delta \mathbf{r}_t = \langle \mathbf{x}, \Delta \mathbf{w}_t \rangle + \mathbf{trd}\,\Delta t,$$

or coordinate-wisely

$$\Delta \mathbf{r}_t(T_l) = \langle \mathbf{x}(T_l), \Delta \mathbf{w}_t \rangle + \mathbf{trd}(T_l)\, \Delta t,\ \ l = 1, ..., k.$$

Define $D_{\mathbf{x}_l}$'s in the same manner as (1.2), which should be called *generalized durations*, and define also

$$D^2_{\mathbf{x}^i, \mathbf{x}^j} := (\mathrm{PV})^{-1} \partial_{\mathbf{x}^i} \partial_{\mathbf{x}^j} (\mathrm{PV}) = (\mathrm{PV})^{-1} \sum_{l=1}^{k} \mathbf{x}^i(T_l) \mathbf{x}^j(T_l)(T_l - t)^2 e^{-\mathbf{r}(T_l)(T_l - t)}, \tag{1.8}$$

which would be called *generalized convexities*. In these settings (1.5) turns into

$$\begin{aligned}
\frac{\Delta(\mathrm{PV})}{\mathrm{PV}} &\simeq -\sum_{j=1}^{n} D_{\mathbf{x}^j}(\Delta w_t^j) + \frac{1}{2} \sum_{1 \leq i,j \leq n} D^2_{\mathbf{x}^i, \mathbf{x}^j} (\Delta w_t^i \Delta w_t^j) \\
&\qquad\qquad + \left\{ (\mathrm{PV})^{-1} \partial_t (\mathrm{PV}) - D_{\mathbf{trd}} \right\} \Delta t \\
&\text{(in the vector notation with } \mathbf{D}_{\mathbf{x}} = (D^1_{\mathbf{x}}, ..., D^n_{\mathbf{x}})) \\
&= -\langle \mathbf{D}_{\mathbf{x}}, \Delta \mathbf{w}_t \rangle + \frac{1}{2PV_t} \langle \mathbf{D}_{\mathbf{x}} \otimes \mathbf{D}_{\mathbf{x}}, \Delta \mathbf{w}_t \otimes \Delta \mathbf{w}_t \rangle \\
&\qquad\qquad + \left\{ (\mathrm{PV})^{-1} \partial_t (\mathrm{PV}) - D_{\mathbf{trd}} \right\} \Delta t \\
&\text{(in this notation convexities are not needed!).}
\end{aligned} \tag{1.9}$$

The parameters $\mathbf{x}^j$'s (including the number of factors n) can be determined, for example, by the *principal component analysis* (PCA) (see [20] or [6, Section 6.4], and in the present paper a brief survey of PCA is given in Appendix A.1.). It should be noted, however, that such a statistical estimation method is founded on the basis of **stationarity**: one needs plenty of homogeneous sample data.

1.3 Stationary models

T.S.Y. Ho, and S. B. Lee [13] proposed such an arbitrage-free *stationary* interest rate model. In essence Ho-Lee model assumes that

$$\{ P^{(t)}(T) := -T^{-1} \log \mathbf{r}_t(t + T) : T \in \mathbf{Z}_+ \} \tag{1.10}$$

is a *random walk* in $\mathbf{R}^{\mathbf{Z}_+}$. By a *random walk* we mean a sum of i.i.d. random variables. In other words, a random walk is a process with stationary independent increments, and hence is consistent with PCA. Actually in Ho-Lee

model increments of the random walk is coin-tossing, and hence is described by a binomial *recombining tree* (see Fig 1 and Fig 2 in section 2.1).

It is widely recognized, however, that the Ho-Lee model is insufficient in that it can describe the movements of "parallel shift" only (see e.g. the quotation in the beginning of section 2.1). In sections 2.1 and 2.2 we will study this widely spread belief in a careful way and show that this is not because the model is one-factor but in fact the problem lies in its parameterization of (1.10) (Theorem 4).

In section 2.3, we will reveal that this puzzle is resolved by shifting to *forward rates*[2]. To exclude trivial arbitrage opportunities continuously compound forward rate over $(T, T']$ at time t should be given by

$$F_t(T, T') := \frac{\mathbf{r}_t(T')(T'-t) - \mathbf{r}_t(T)(T-t)}{T'-T}. \tag{1.11}$$

In this shift it is assumed that for every T, the increments $\Delta F_t(T) := F_t(T) - F_{t-\Delta t}(T)$ of the instantaneous forward rate $F_t(T) := F_t(T, T+\Delta t)$ are i.i.d.., or more precisely, we assume that for a sequence of i.i.d. random variables $\Delta \mathbf{w}$ satisfying (1.7), there exists $\sigma(T) \in \mathbf{R}^n$ and $\mu(T) \in \mathbf{R}$ for each T such that for arbitrary $0 \leq t \leq T$

$$\Delta F_t(T) = \langle \sigma(T), \Delta \mathbf{w}_t \rangle + \mu(T)\Delta t. \tag{1.12}$$

The stochastic processes $F_t(T)$'s are **stationary** in the sense $\sigma(T)$ and $\mu(T)$ are constant with respect to the time parameter t.

In sharp contrast with the cases of Ho-Lee and its direct generalizations there always exists an arbitrage-free stationary model for given term structure of volatilities $\sigma(\cdot) : \mathbf{Z}_+\Delta t \to \mathbf{R}$. In fact, we have the following.

Theorem 1. *For any given (estimated) term structure of volatilities $\sigma(\cdot)$: $\mathbf{Z}_+\Delta t \to \mathbf{R}$ there always exists an arbitrage-free* **stationary** *forward rate model given by (1.12). Precisely speaking, the following holds. The model with (1.12) is arbitrage-free if*

$$\mu(T) = \frac{1}{\Delta t} \log \frac{\mathbf{E}[\exp \langle \rho(T), \Delta \mathbf{w} \rangle]}{\mathbf{E}[\exp \langle \rho(T+\Delta t), \Delta \mathbf{w} \rangle]}, \tag{1.13}$$

where $\rho(T) = \Delta t \sum_{0 \leq u \leq T} \sigma(u)$.

[2] A forward rate is a pre-agreed rate for borrowing during a pre-agreed future time interval.

A proof of Theorem 1 will be given in section 2.3.

Though the shift to forward rate model parallels with the celebrated works by D. Heath, R.A. Jarrow and A. Morton [11] and [12], our model presented here is an innovation in that

(i) The model is stationary, and hence consistent with PCA. (It is not so clear whether the drift μ can be constant in time under no-arbitrage hypothesis.)

(ii) It is a generalization of Ho-Lee's binomial model in the sense that it is within a discrete time-state framework and it can be described by a **recombining tree**; If Δw^j's have only finite possibilities (in fact we will construct under the assumption of $(a2')$ in section 2.2.) then the tree (=all the scenarios) of $\mathbf{w} \equiv (w^1, ..., w^n)$, hence the tree of $\mathbf{r}$, becomes **recombining** since $\Delta\mathbf{w}(t)(w)+\Delta\mathbf{w}(t')(w')$ equals $\Delta\mathbf{w}(t')(w)+\Delta\mathbf{w}(t)(w')$ for any given time t, t' and one step scenarios w, w'.

(iii) It is a **multi-factor** model; along the line of (1.9) and (1.8) it gives an alternative sensitivity analysis beyond *duration* and *convexity* (see section 3).

(iv) Furthermore, this class can be seen as a discrete-time analogue of (a special cases of) multi-factor generalization of Ritchken-Sankarasubramanian's model [24] due to K. Inui and M. Kijima [17] (see section 4).

2 Linear models

2.1 Ho-Lee model revisited

> The Ho-Lee model was a significant improvement over what came before, but it had a number of failings. It was a one-factor model, and the way the term structure evolved over time was through parallel shifts. (R. Jarrow [19])

Let $P^{(t)}(T)$ denote the price at time t of zero-coupon bond whose time to maturity is T: [3] To avoid confusions, below we give a reversed form of (1.10):

[3]Note that this notation is different from the standard one. The parameter T usually stands for the maturity date as we have done and will do for $\mathbf{r}$ and others. The parameterization used here originates from [13].

$$\mathbf{r}_t(T) = \frac{1}{T-t}\log P^{(t)}(T-t), \quad (T>t). \tag{2.1}$$

Here t and T are non-negative integers; in particular, Δt is set to be 1. Note that P must be positive at least and $P^{(t)}(0) = 1$ without fail.

The Ho-Lee model is characterized by the following assumptions $(a1)$ to $(a4)$.

$(a1)$ The random variable $P^{(t)}(T)$ depends only on so-called *state variable* $i = i_t$ [(A4) of pp1013 in [13]].

$(a2)$ For each t, $\Delta i_t = i_{t+1} - i_t$ can take only two values; as in [13] the two are 1 and 0 [(4) of pp1014 in [13]].

As a consequence of $(a2)$, the state space of i is $\mathbf{Z}$, and if we take $i_0 = 0$, then it can be $\mathbf{Z}_+$. We sometimes write $P_i^{(t)}(T)$, and in this case we think of P as a function on $\mathbf{Z}_+(\text{time}) \times \mathbf{Z}_+(\text{state}) \times \mathbf{Z}_+(\text{maturity})$.

The *tree* of the assumption (a2) is illustrated in Fig 1. The first assump-

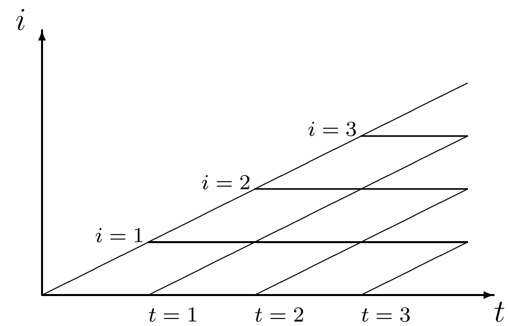

Figure 1: Recombining tree #1

tion (a1) also says that the function (often referred to as *term structure*) $P^{(t)}(\cdot) : \mathbf{Z}_+ \to \mathbf{R}$, or equivalently the random vector $(P^{(t)}(1), P^{(t)}(2), ..., P^{(t)}(T), ...)$ depends only on i_t. In other words, a realization of term structure is attached to each node of the tree of Fig 1 (See Fig 2).

The next assumption $(a3)$ makes Ho-Lee model as it is [(7), (8), and (10) of pp1017 in [13]].

$(a3)$ The model is *stationary* in the sense that $P_{i_t}^{(t)}(T)P_{i_{t-1}}^{(t-1)}(1)/P_{i_{t-1}}^{(t-1)}(T+1)$ depends only on T and $\Delta i_t = i_t - i_{t-1}$.

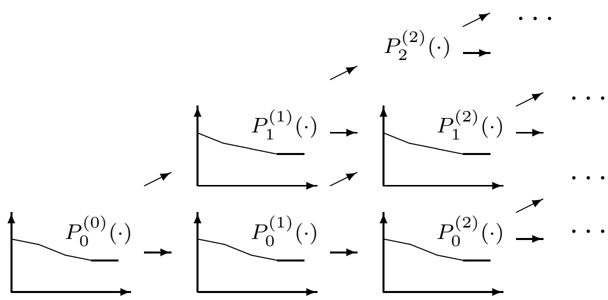

Figure 2: Term structures attached to the recombining tree

The assumption $(a3)$ says that if $\{\Delta i_t : t \in \mathbf{N}\}$ are i.i.d., so are $\{P_{i_t}^{(t)}(\cdot)P_{i_{t-1}}^{(t-1)}(1) / P_{i_{t-1}}^{(t-1)}(\cdot + 1) : (t \in \mathbf{N})\}$. In this sense we call it stationary. Actually we implicitly assume that under the real world measure Δi_t are i.i.d.. Denote

$$H(T, \Delta i_t) := \frac{P_{i_t}^{(t)}(T)P_{i_{t-1}}^{(t-1)}(1)}{P_{i_{t-1}}^{(t-1)}(T+1)}. \tag{2.2}$$

Recall that $h(T) := H(T, 1)$ and $h^*(T) := H(T, 0)$ are called perturbation function(s) in [13].

The next condition $(a4)$ says that the model is arbitrage-free (see Theorem 18 in Appendix).

$(a4)$ There exists $\pi \in (0, 1)$ such that for arbitrary $t, i \in \mathbf{Z}_+$ and $T > 0$,

$$P_i^{(t)}(T+1) = \pi P_i^{(t)}(1)P_{i+1}^{(t+1)}(T) + (1-\pi)P_i^{(t)}(1)P_i^{(t+1)}(T).$$

Under these assumptions, dynamics of the entire term structure is fully determined. More precisely, we have the following.

Theorem 2. *(i) Under the assumption* $(a1)$, $(a2)$, *and* $(a3)$, $\{P^{(t)}(\cdot) : t \geq t_0\}$ *as* $\mathbf{R}^{\mathbf{Z}_+}$ *valued stochastic process is parameterized by given initial term structure* $\mathbf{r}_{t_0}$, *the constant volatility* σ, *and the drift function* $\mu : \mathbf{Z}_+ \to \mathbf{Z}_+$ *as*

$$P_i^{(t)}(T) = \frac{P^{(0)}(t+T)}{P^{(0)}(t)} \exp\{-T\sigma(2i-t) - \sum_{u=0}^{T-1}\sum_{v=1}^{t} \mu(u+v-1)\}. \tag{2.3}$$

(ii) The no-arbitrage condition (a4) is equivalent to the following "drift condition": for a $\pi \in (0,1)$ we have

$$\mu(T) = \log \frac{\pi e^{-(T+1)\sigma} + (1-\pi)e^{(T+1)\sigma}}{\pi e^{-T\sigma} + (1-\pi)e^{T\sigma}}. \tag{2.4}$$

This theorem will be proved as a special case ($n = 2$ and $i = 2w - 1$) of Theorem 4 below.

Corollary 3 (Ho and Lee [13], (19) and (20) of pp 1019). *Under (a1), (a2), and (a3), no-arbitrage condition (a4) is equivalent to the existence of $\pi \in (0,1)$ and $\delta > 0$ such that*

$$h(T) = \frac{1}{\pi + (1-\pi)\delta^T} \text{ and } h^*(T) = \frac{\delta^T}{\pi + (1-\pi)\delta^T} \tag{2.5}$$

for all $T \in \mathbf{Z}_+$.

Proof. By (2.13) below we have

$$\begin{aligned} H(T, \Delta i) &= e^{-T\sigma(2\Delta i - 1)} e^{-\sum_{u=0}^{T-1} \mu(u)}, \\ &\quad \text{(by substituting (2.4))} \\ &= e^{-\sigma T(2\Delta i - 1)} \prod_{u=0}^{T-1} \frac{\pi e^{-u\sigma} + (1-\pi)e^{u\sigma}}{\pi e^{-(u+1)\sigma} + (1-\pi)e^{(u+1)\sigma}} \\ &= \frac{e^{-\sigma T(2\Delta i - 1)}}{\pi e^{-T\sigma} + (1-\pi)e^{T\sigma}}. \end{aligned}$$

Therefore, we have

$$h(T) = H(T, 1) = \frac{e^{-\sigma T}}{\pi e^{-T\sigma} + (1-\pi)e^{T\sigma}} = \frac{1}{\pi + (1-\pi)e^{2T\sigma}}$$

and

$$h^*(T) = H(T, 0) = \frac{e^{\sigma T}}{\pi e^{-T\sigma} + (1-\pi)e^{T\sigma}} = \frac{e^{2\sigma T}}{\pi + (1-\pi)e^{2T\sigma}}.$$

By putting $\delta = e^{2\sigma}$ we obtain (2.5). $\square$

2.2 Multi-dimensional generalization

As we have discussed in section 1.2, by a multi-factor generalization we want to mean such as (1.6) and (1.12), with the moment condition (1.7).

For the time being let us consider the condition (1.7). Suppose that $\Delta\mathbf{w}$ is defined on a finite set, say, $S = \{0, 1, ..., s\}$. Let $L(S) := \{f : S \to \mathbf{R}\}$. We endow the scholar product with $L(S)$ by $\langle x, y\rangle = \mathbf{E}[xy]$. Since $\Delta t = 1$ here in this section, the condition (1.7) is rephrased as:

$$\{1, \Delta w^1, ..., \Delta w^n\} \text{ is an orthonormal system of } L(S).$$

(For general Δt, it is replaced with $\{1, \Delta w^1/\sqrt{\Delta t}, ..., \Delta w^n/\sqrt{\Delta t}\}$.) From this observation we notice that $s+1 = \sharp S$ should be greater than n. If $s > n$, then we can extend $\{1, \Delta w^1, ..., \Delta w^n\}$ to have an orthonormal basis of $L(S)$ by constructing (dummy) random variables $\Delta w^{n+1}, ..., \Delta w^s$. Thus for simplicity we want to assume $n = s$.

Let us go a little further. Let

$$m_{i,j} = \Delta w^i(j)\sqrt{\Pr(\{s_j\})}, \quad 0 \le i \le n,\, 0 \le j \le s, \tag{2.6}$$

where $\Delta w^0 \equiv 1$. When $n = s$, then $M = (m_{i,j})_{0\le i,j\le n=s}$ is an orthogonal matrix; following a standard notation, we have $M \in O(n+1)$. Conversely, given an orthogonal matrix $M = (m_{i,j})_{0\le i,j\le n} \in O(n+1)$ with $m_{0,j} > 0$ for all j, we can construct a random variable $\Delta\mathbf{w} = (\Delta w^1, ..., \Delta w^n)$ satisfying (1.7) by setting $m_{0,j} = \sqrt{\Pr(\{s_j\})}$ and the relation (2.6).

With the above considerations in mind, we generalize the assumption $(a2)$ as follows.

$(a2')$ The state variable i_t is $\mathbf{R}^{n+1}$-valued, and $\Delta i_t = (\Delta t, \Delta w_t^1, ..., \Delta w_t^n)$, where $(\Delta w_t^1,\ ..., \Delta w_t^n)$, $t = 1, 2, ...$ are independent copies of $\Delta\mathbf{w}$; a random variable satisfying (1.7) and taking only $n+1$ distinct values.

Note that $(a2) + (a1)$ can be seen as a special case of $(a2') + (a1)$ by setting $n = 1$, $\Delta w^1(0) = -1, \Delta w^1(1) = 1$, and $\Pr(\{0\}) = \Pr(\{1\}) = 1/2$ In fact, $\hat{i}_t = w_t^1/2 + t/2$, is a function of $i_t = (t, w_t)$, and $\Delta\hat{i}_t$ can be either 1 or 0. Conversely, starting from $(a2)$, the function $(1, 2i-1)$ on S can be orthonormal basis of $L(S)$.

We denote the state space of $\mathbf{w}_t$ by $\mathfrak{Z}$. In general the state space of $i_t = (t, \mathbf{w})$ is described by a recombining tree, the number of whose nodes

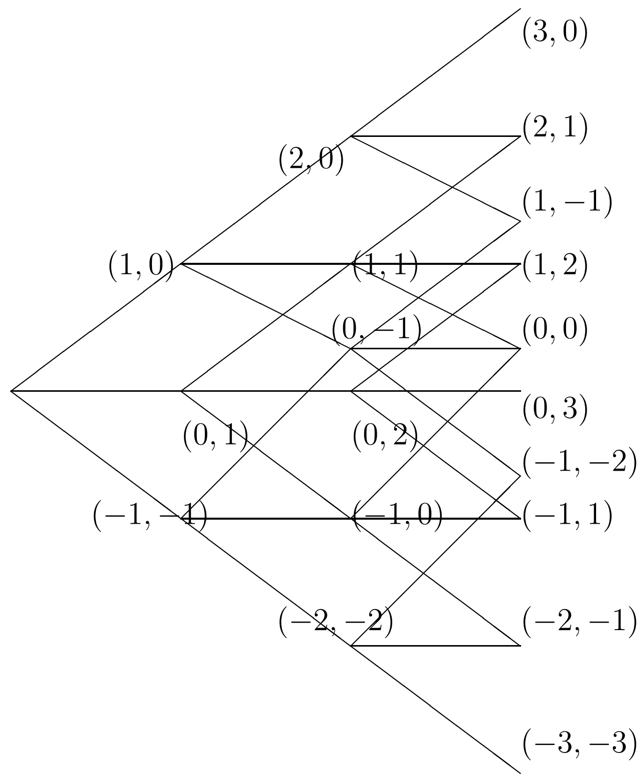

Figure 3: Recombining tree #2. (Note that the pairs of numbers does not mean the values of w.)

grow in an order of t^n. Detailed studies of the construction of $\mathbf{w}$ and its tree are presented in [2]. Here we illustrate a tree for $n = 2$ in Fig 3.

According as this generalization, $(a4)$ also needs to be modified.

$(a4')$ There exists a probability (measure) $\pi_i(\cdot) \in (0, 1)$ on S such that for arbitrary $T > 0$ and $i = (t, w) \in \mathfrak{Z}$,

$$P_i^{(t)}(T+1) = P_i^{(t)}(1) \sum \pi_i(s) P_{i+\Delta i(s)}^{(t+1)}(T).$$

As we have remarked in section 1.3, the stationarity assumption $(a3)$ is yet too restrictive even under these relaxed conditions.

Theorem 4. *(i) Under the assumption* $(a1)$, $(a2')$, *and* $(a3)$, *dynamics of the entire term structure* $\{P^{(t)}(\cdot) : t \geq 0\}$ *as* $\mathbf{R}^{\mathbf{Z}_+}$ *valued stochastic process*

is parameterized by given initial term structure $P^{(0)}(\cdot) : \mathbf{R}^{\mathbf{Z}_+}$, the constant volatility $\sigma \in \mathbf{R}^n$, and the drift function $\mu : \mathbf{Z}_+ \to \mathbf{R}$ as

$$P^{(t)}(T) = \frac{P^{(0)}(t+T)}{P^{(0)}(t)} \exp\{-T\langle\sigma, \mathbf{w}_t - \mathbf{w}_0\rangle - \sum_{u=0}^{T-1}\sum_{v=1}^{t}\mu(u+v-1)\}. \quad (2.7)$$

(ii) The no-arbitrage condition $(a4')$ is equivalent to the following "drift condition": for a probability π on S we have

$$\mu(T) = \log \frac{\sum_s \pi(s)e^{-(T+1)\langle\sigma,\Delta\mathbf{w}(s)\rangle}}{\sum_s \pi(s)e^{-T\langle\sigma,\Delta\mathbf{w}(s)\rangle}}. \quad (2.8)$$

Proof. (i) First, notice that the assumption $(a1)$ in terms of $\mathbf{f}_t(T) = F_t(t+T, t+T+\Delta t)$, the forward rate over $(T-t, T+1-t]$ at time t, is stable: If $(a1)$ holds, then $\mathbf{f}_t(T)$ depends only on i_t and vise versa. To see this, just use the following expressions:

$$\mathbf{f}_t(T) = \log \frac{P^{(t)}(T)}{P^{(t)}(T+1)},$$

and

$$P^{(t)}(T) = \exp\{-\sum_{u=0}^{T-1}\mathbf{f}_t(u)\}. \quad (2.9)$$

Next, observe that the assumption $(a3)$ in terms of $\mathbf{f}_t(T)$ turns into the following equation:

$$\mathbf{f}_t(T) - \mathbf{f}_{t-1}(T+1) = \log H(T, \Delta i_t) - \log H(T+1, \Delta i_t). \quad (2.10)$$

Thanks to the "completeness" assumption, the right-hand-side of (2.10) is reparametrized as $\langle\sigma(T), \Delta i_t\rangle_{\mathbf{R}^n}$ by some $\sigma(T) \in \mathbf{R}^{n+1}$ for each $T \in \mathbf{Z}_+$. Note that the choice of $\sigma(T)$ is unique. With this representaion, the solution to (2.10) is obtained as

$$\mathbf{f}_t(T) = \mathbf{f}_0(T+t) + \sum_{u=1}^{t}\langle\sigma(T+t-u), \Delta i_u\rangle. \quad (2.11)$$

Since i_t is an exchangeable in the sense that

$$i(s_1, ..., s_t) = i(s_{\tau(1)}, ..., s_{\tau(t)})$$

for arbitrary permutation $\tau \in \mathfrak{S}_t$, so is $\mathbf{f}_t(T) \equiv \mathbf{f}_t(T; s_1, ..., s_t)$, $(s_1, ..., s_t) \in S^t$. Therefore, from the explicit expression (2.11) we have

$$\begin{aligned}\sum_{u=1}^{t} \langle \sigma(T+t-u), \Delta i(s_u) \rangle &= \sum_{u=1}^{t} \langle \sigma(T+t-u), \Delta i(s_{\tau(u)}) \rangle \\ &= \sum_{u=1}^{t} \langle \sigma(T+t-\tau^{-1}(u)), \Delta i(s_u) \rangle\end{aligned}$$

for arbitrary permutation τ. In particular, the relation holds for any transposition. Therefore, we have

$$\begin{aligned}&\langle \sigma(T+t-u) - \sigma(T+t-u'), \Delta i(s_u) - \Delta i(s_{u'}) \rangle_{\mathbf{R}^{n+1}} \\ &\quad \text{(since the first coordinate of } i \text{ is independent of } s) \\ &\quad = \langle \sigma'(T+t-u) - \sigma'(T+t-u'), \Delta \mathbf{w}(s_u) - \Delta \mathbf{w}(s_{u'}) \rangle_{\mathbf{R}^n},\end{aligned}$$

where σ' is the n-dimensional vector projected from the latter n components of σ.

On the other hand, since $M = (m_{i,j})_{0 \le i,j \le n}$ is regular, so is (at least) one of the $n \times n$ matrices $M_k := (m_{i,j})_{1 \le i \le n, 0 \le j \le n, j \ne k}$, $k = 0, 1, ..., n$. This is because $\det M = \sum_k m_{0,k}(-1)^{n+k} \det M_k$. Consequently, setting $s_{u'} = k$, we have that $\{\Delta \mathbf{w}(j) - \Delta \mathbf{w}(k),\ j = 0, 1, ..., n,\ j \ne k\}$ span the whole space $\mathbf{R}^n$. Therefore, we have

$$\sigma'(T+t-u) - \sigma'(T+t-k) = 0 \text{ for arbitrary } 1 \le u \le t. \tag{2.12}$$

Since (2.12) is true for any $T \in \mathbf{Z}_+$, $\sigma'(T)$ equals to a constant $\sigma \in \mathbf{R}^n$ irrespective of T.

Reminding that first component of σ can be taken arbitrary, which we denote by μ, we have

$$\mathbf{f}_t(T) = \mathbf{f}_0(T+t) + \langle \sigma, \mathbf{w}_t - \mathbf{w}_0 \rangle + \sum_{u=1}^{t} \mu(T+u-1).$$

Hence, by (2.9) we have (2.7).

(ii) Back to the original parameterization, we have seen

$$\log H(T, \Delta i_t) - \log H(T+1, \Delta i_t) = \langle \sigma, \Delta \mathbf{w}_t \rangle + \mu(T), \quad T \in \mathbf{Z}_+.$$

Therefore, (since $H(0,\cdot) = P(0) = 1$)

$$\log H(T, \Delta i) = -T\langle \sigma, \Delta \mathbf{w}_t \rangle - \sum_{u=0}^{T-1} \mu(u). \tag{2.13}$$

On the other hand, since the assumption $(a4')$ in terms of H is

$$1 = \sum_s \pi_i(s) H(T, \Delta i(s)),$$

we have

$$e^{\sum_{u=1}^{T-1} \mu(u)} = \sum_s \pi_i(s) e^{-T\langle \sigma, \Delta \mathbf{w}_t(s) \rangle},$$

and hence (2.8). □

Remark 5. Dynamics of term structure in terms of spot rates is now

$$\Delta \mathbf{r}(\cdot) = \langle \sigma, \Delta \mathbf{w} \rangle + (\text{deterministic}).$$

This means that under $(a3)$ Ho-Lee model is still poor; it can describe only parallel shifts. In this sense this is *not* a multi-factor generalization.

Remark 6. (This was suggested by Damir Filipovic in a private communication) The equation (2.10), or reparametrized one

$$\mathbf{f}_t(T) - \mathbf{f}_{t-\Delta t}(T + \Delta t) = \langle \sigma, \Delta \mathbf{w}_t \rangle + \mu(T), \tag{2.14}$$

is a discrete-time analogue of a class of Stochastic Partial Differential Equations (SPDE) studied by Brace, Gatarek, and Musiala [4]. In fact, at least symbolically (2.14) can be rearranged into

$$\Delta \mathbf{f}_t(T) = (\partial_T \mathbf{f}_{t-\Delta t} + \mu'(T))\Delta t + \langle \sigma, \Delta \mathbf{w}_t \rangle,$$

where $\mu'(T) = \mu(T)/\Delta t$,

$$\Delta \mathbf{f}_t(T) \equiv \mathbf{f}_t(T) - \mathbf{f}_{t-\Delta t}(T)$$

and

$$\partial_T \mathbf{f}_{t-\Delta t} \equiv \frac{1}{\Delta t}\{\mathbf{f}_{t-\Delta t}(T + \Delta t) - \mathbf{f}_{t-\Delta t}(T)\}.$$

2.3 Multi-factor generalization

> Instead of focusing upon bond prices as in Ho and Lee (1986), we concentrate on forward rates. This "modification" makes the model easier to understand and to perform mathematical analysis. (D. Heath, R. Jarrow and A. Morton [11])

We have shown in the previous section that $(a3)$ is too restrictive to have rich models. One may well ask if there would be another stationary assumption other than $(a3)$. The simplest alternative candidate may be

$$-\frac{1}{T}\log P^{(t)}(T) = \langle \sigma(T), \mathbf{w}_t \rangle + \mu(T)\, t, \tag{2.15}$$

but:

Proposition 7. *The model (2.15) cannot be arbitrage-free unless $\sigma(T) = \sigma(1)$ for arbitrary $T \geq 1$. Namely only parallel shifts are allowed for.*

Proof. By Theorem 18 (iv) we know that a model is arbitrage free if and only if

$$\mathbf{E}^{\pi}[P^{(t+\Delta t)}(T)P^{(t)}(1)|\mathcal{F}_t] = P^{(t)}(T+1)$$

with a probability π on $\mathcal{F}_t$. Applying this to (2.15), we have

$$\begin{aligned} &\mathbf{E}^{\pi}[e^{\langle \sigma(T), \Delta \mathbf{w}\rangle + \mu(T)\Delta t}] \\ &= \exp\{\langle (T+1)\sigma(T+1) - T\sigma(T) - \sigma(1), \mathbf{w}_t\rangle \\ &\qquad\qquad + (T+1)\mu(T+1) - T\mu(T) - \mu(1)\}. \end{aligned}$$

While the left-hand-side is a constant, the right-hand-side is not unless

$$(T+1)\sigma(T+1) - T\sigma(T) - \sigma(1) = 0.$$

This means $\sigma(T) \equiv \sigma(1)$ for arbitary T. □

To summarize, it would be safe to say that stationarity requirements are difficult to be compatible with no-arbitrage restrictions. As we have mentioned in section 1, however, we can resolve the problem by changing the parameterization of maturity; assuming (1.12) instead of $(a3)$.

Let us introduce (or come back to) the following (standard) notation:

$$P_t^T = P^{(t)}(t+T) \quad t \leq T.$$

(in this case T means "maturity", not "time to maturity"). Then

$$F_t(T) = \frac{1}{\Delta t} \log \frac{P_t^T}{P_t^{T+\Delta t}}. \tag{2.16}$$

and

$$P_t^T = \exp\{-\Delta t \sum_{u=t}^{T-\Delta t} F_t(u)\}. \tag{2.17}$$

Here and hereafter the unit Δt is not necessarily 1.

By an *interest rate model* we mean a family of strictly positive stochastic process $\{P_t^T\}$, which are, at each time t, a function of $(\Delta \mathbf{w}_1, ..., \Delta \mathbf{w}_t)$. In a more sophisticated way of speaking, they are adapted to the natural filtration of the random walk $\mathbf{w}$. We will denote $\mathcal{F}_t := \sigma(\Delta \mathbf{w}_1, ..., \Delta \mathbf{w}_t)$. Then $\{\mathcal{F}_t\}$ is the natural filtration.

To prove Theorem 1 we rely on the following theorem, which is actually a corollary of Theorem 18.

Theorem 8. *An interest rate model $\{P_t^T\}$ is arbitrage-free if and only if there exists a strictly positive $\{\mathcal{F}_t\}$-adapted stochastic process D_t such that*

$$P_t^T = \mathbf{E}[\mathsf{D}_T \,|\, \mathcal{F}_t]/\mathsf{D}_t \tag{2.18}$$

holds for arbitrary $t \leq T$.

A proof of Theorem 8 will be given in Appendix A.2. As a corollary we have the following *consistency condition with respect to the initial forward rate curve* for interest rate modeling.

Corollary 9. *An interest rate model $\{P_t^T\}$ is arbitrage-free if and only if there exists a strictly positive $\{\mathcal{F}_t\}$-adapted stochastic process $\widehat{\mathsf{D}}_t$ with $\mathbf{E}[\widehat{\mathsf{D}}_t] = 1$ for all $t \geq 0$ such that (i)*

$$P_t^T = \frac{P_0^T \mathbf{E}[\widehat{\mathsf{D}}_T \,|\, \mathcal{F}_t]}{P_0^t \widehat{\mathsf{D}}_t}$$

holds for arbitrary $t \leq T$, or equivalently (ii) (in terms of the instantaneous forward rate)

$$F_t(T) = F_0(T) + \frac{1}{\Delta t} \log \frac{\mathbf{E}[\widehat{\mathsf{D}}_T \,|\, \mathcal{F}_t]}{\mathbf{E}[\widehat{\mathsf{D}}_{T+\Delta t} \,|\, \mathcal{F}_t]}$$

holds for arbitrary $t \leq T$.

Proof. Observing $P_0^T = \mathbf{E}[\mathsf{D}_T]/\mathsf{D}_0$ by setting $t = 0$ in (2.18), the first statement (i) can be implied from (2.18) by putting $\widehat{\mathsf{D}}_t = \mathsf{D}_t/\mathbf{E}[\mathsf{D}_t]$. The converse, from (i) to (2.18), is trivial. The equivalence of (i) and (ii) is obtained from (2.17) and (2.16). $\square$

Now we are in a position to prove Theorem 1.

Proof of Theorem 1. Let $\mathsf{D}_t := \exp\langle \rho(t), \mathbf{w}_t\rangle$. Then

$$\begin{aligned}
\mathbf{E}[\mathsf{D}_T|\mathcal{F}_t] &= \mathbf{E}[\exp\langle \rho(T), \mathbf{w}_T\rangle|\mathcal{F}_t] \\
&= \exp\langle \rho(T), \mathbf{w}_t\rangle \mathbf{E}[\exp\langle \rho(T), \mathbf{w}_T\rangle] \\
&= \exp\langle \rho(T), \mathbf{w}_t\rangle \prod_{u=t+\Delta t}^{T} \mathbf{E}[\exp\langle \rho(T), \Delta\mathbf{w}_u\rangle] \\
&= \exp\langle \rho(T), \mathbf{w}_t\rangle (\mathbf{E}[\exp\langle \rho(T), \Delta\mathbf{w}\rangle])^{(T-t)/\Delta t}.
\end{aligned}$$

In particular we have

$$\mathbf{E}[\mathsf{D}_t] = \exp\langle \rho(t), \mathbf{w}_0\rangle (\mathbf{E}[\exp\langle \rho(t), \Delta\mathbf{w}\rangle])^{t/\Delta t}.$$

Let

$$\widehat{\mathsf{D}}_t := \mathsf{D}_t/\mathbf{E}[\mathsf{D}_t] = \exp\langle \rho(t), \mathbf{w}_t - \mathbf{w}_0\rangle (\mathbf{E}[\exp\langle \rho(T), \Delta\mathbf{w}\rangle])^{-t/\Delta t}.$$

Then by Theorem 8 the model with

$$\begin{aligned}
F_t(T) &:= F_0(T) + \frac{1}{\Delta t}\log\frac{\mathbf{E}[\widehat{\mathsf{D}}_T \,|\, \mathcal{F}_t]}{\mathbf{E}[\widehat{\mathsf{D}}_{T+\Delta t} \,|\, \mathcal{F}_t]} \\
&= F_0(T) + \frac{1}{\Delta t}\langle \rho(T) - \rho(T+\Delta t), \mathbf{w}_t - \mathbf{w}_0\rangle \\
&\quad + \frac{T-t}{(\Delta t)^2}\log\mathbf{E}[\exp\langle \rho(T), \Delta\mathbf{w}\rangle] \\
&\quad - \frac{T+\Delta t - t}{(\Delta t)^2}\log\mathbf{E}[\exp\langle \rho(T+\Delta t), \Delta\mathbf{w}\rangle],
\end{aligned} \tag{2.19}$$

is arbitrage-free. Noting that $\rho(T) - \rho(T+\Delta t) = \sigma(T)\Delta t$, from (2.19) we can deduce

$$\Delta F_t(T) = \langle \sigma(T), \Delta\mathbf{w}_t\rangle + \frac{1}{\Delta t}\log\frac{\mathbf{E}[\exp\langle \rho(T), \Delta\mathbf{w}\rangle]}{\mathbf{E}[\exp\langle \rho(T+\Delta t), \Delta\mathbf{w}\rangle]}.$$

Thus we have proved the assertion. $\square$

3 Application to sensitivity analysis

3.1 parameter estimation

Let us come back to the sensitivity analysis. To fit reality, we assume that we can observe only such forward rates as $F_t(T_i, T_{i+1})$ (for past t, of course) for a coarse set of maturities $\{T_1, ..., T_k\}$; it may happen that $T_i - T_{i+1} \gg \Delta t$.

It should be noted that from (1.11) we have

$$(T' - T)F_t(T, T') + (T'' - T')F_t(T', T'') = (T'' - T)F_t(T, T'') \tag{3.1}$$

for arbitrary $t < T < T' < T''$. Therefore, the assumption (1.12) implies the following:

$$\begin{aligned} \Delta F_t(T_i, T_{i+1}) &:= F_{t+\Delta t}(T_i, T_{i+1}) - F_t(T_i, T_{i+1}) \\ &= \langle \sigma(T_i, T_{i+1}), \Delta \mathbf{w}_t \rangle + \mu(T_i, T_{i+1})\Delta t, \end{aligned} \tag{3.2}$$

where

$$\sigma_j(T_i, T_{i+1}) = \frac{\Delta t}{T_{i+1} - T_i} \sum_{u=T_i}^{T_{i+1}-\Delta t} \sigma_j(u) = \frac{\rho(T_{i+1}) - \rho(T_i)}{T_{i+1} - T_i}, \tag{3.3}$$

for $j = 0, 1, ..., n$, with the convention that $\sigma_0 = \mu$. The drift condition (1.13) implies

$$\mu(T_i, T_{i+1}) = \frac{1}{T_{i+1} - T_i} \log \frac{\mathbf{E}[\exp \langle \rho(T_i), \Delta \mathbf{w} \rangle]}{\mathbf{E}[\exp \langle \rho(T_{i+1}), \Delta \mathbf{w} \rangle]}. \tag{3.4}$$

Note that for every i, $\Delta F_t(T_i, T_{i+1})$ are again i.i.d. for all $0 \le t \le T_i$.

In practice, up to time t ($< T_1$) we may obtain sufficiently many (though it is less that $t/\Delta t$) homogeneous sample data of $\Delta F(T_i, T_{i+1})$ for $i = 1, ..., k-1$, and then by PCA (see A.1), we can estimate (together with n itself) the *volatility matrix* $(\sigma_{i,j})$ with $\sigma_{i,j} = \sigma_j(T_i, T_{i+1}) \in \mathbf{R}^k \otimes \mathbf{R}^n$.

To obtain a full term structure of volatilities $\sigma : \mathbf{Z}_+\Delta t \to \mathbf{R}$ one needs to interpolate the volatility matrix σ. An easiest interpolation is given by

$$\sigma^j(T) = \sigma_{ij} \equiv \sigma^j(T_i, T_{i+1}) \text{ if } T_i < T \le T_{i+1}, i = 0, 1, ..., k, \tag{3.5}$$

where $j = 0, 1, ..., n$ with conventions of $T_0 = -\infty$, $T_k = +\infty$, and $\sigma_{i0} = \sigma^0(T_i, T_{i+1}) = \mu(T_i, T_{i+1})$.

Remark 10. It may happen that a statistical estimator of μ (possibly by sample average) does not necessarily satisfy the drift condition. Note, however, that neither (1.13) nor (3.4) is if and only if condition; there might be other arbitrage-free models that are consistent with estimated volatility matrix. A careful study on the existence of arbitrage-free model that is consistent with both estimated volatility and average [4] is still open and would be interesting, both theoretically and practically.

Remark 11. It should be noted that in our model, we first work on the real world measure, and then estimate the risk neutral measure (see Theorem 18 (iv)). The stationarity cannot expect in general under the risk neutral measure.

3.2 A new framework of sensitivity analysis

Based on the stationary model by Theorem 1 and the estimated parameters we now introduce a new framework of sensitivity analysis.

Theorem 12. *We have the following random expansion with respect to* Δt *of the present value* PV *of a given cash flow* $\mathbf{CF}$*:*

$$\begin{aligned}\frac{\Delta(\mathrm{PV})}{\mathrm{PV}} \simeq \sum_l \mathbf{CF}(T_l) e^{-(T_l - t)\mathbf{r}_t(T_l)} \Bigg\{ &\langle \{\rho(T_l) - \rho(t)\}/\mathrm{PV}_t, \Delta \mathbf{w}_t \rangle \\ &+ \frac{1}{2\mathrm{PV}_t} \langle \{\rho(T_l) - \rho(t)\} \otimes \{\rho(T_l) - \rho(t)\}, \Delta \mathbf{w}_t \otimes \Delta \mathbf{w}_t \rangle \\ &+ \Big(\langle \sigma(T_1, T_2), \mathbf{w}_t - \mathbf{w}_0 \rangle - \{\rho^0(T_l) - \rho^0(t)\} + \rho^0(t)t \Big) \Delta t \Bigg\} + \mathrm{o}(\Delta t).\end{aligned} \tag{3.6}$$

Here by $\rho^0(t)$ *we mean* $\sum_{0<u\leq t} \mu(u)$*. Consequently,*

$$\mathrm{PV}^{-1} \sum_l (\rho^j(T_l) - \rho^j(t)) \mathbf{CF}(T_l) e^{-(T_l - t)\mathbf{r}_t(T_l)}, \ j = 1, ..., n \tag{3.7}$$

are generalized durations and

$$\begin{aligned}\mathrm{PV}^{-1} \sum_l (\rho^i(T_l) - \rho^i(t))(\rho^j(T_l) - \rho^j(t)) \mathbf{CF}(T_l) e^{-(T_l - t)\mathbf{r}_t(T_l)}&, \\ 1 \leq i, j \leq n&,\end{aligned} \tag{3.8}$$

[4] This is not the one by drift conditions, of course.

are generalized convexities in the sense of section 1.2.

Note that when $\rho(T)$ is affine in $T \iff \sigma(T) \equiv \sigma$; i.e., only *parallel shifts* are allowed, the standard duration and convexity are retrieved.

Proof. Using (3.5) we can represent the spot rate $\mathbf{r}$ by the estimated volatility matrix (and average):

$$\begin{aligned}
\mathbf{r}_t(T_i) &= -\frac{1}{T_i - t}\log P_t^{T_i} = -\frac{\Delta t}{T_i - t}\sum_{u=t}^{Ti-\Delta t} F_t(u) \\
&= -\frac{T_1 - t}{T_i - t}F_t(t, T_1) - \frac{T_2 - T_1}{T_i - t}F_t(T_1, T_2) - \cdots - \frac{T_i - T_{i-1}}{T_i - t}F_t(T_{i-1}, T_i) \\
&= -\frac{1}{T_i - t}\langle \Delta t \sum_{u=t}^{T_1-\Delta t} \boldsymbol{\sigma}(u) + \sum_{l=1}^{i-1}(T_{l+1} - T_l)\boldsymbol{\sigma}(T_l, T_{l+1}), i_t - i_0\rangle \\
&= -\frac{1}{T_i - t}\langle (T_1 - t)\boldsymbol{\sigma}(T_1, T_2) + \sum_{l=1}^{i-1}(T_{l+1} - T_l)\boldsymbol{\sigma}(T_l, T_{l+1}), i_t - i_0\rangle \\
&\qquad\qquad\qquad\qquad - \frac{1}{T_i - t}\log P_0^{T_i} \\
&= -\frac{1}{T_i - t}\{\langle \boldsymbol{\rho}(T_l) - \boldsymbol{\rho}(t), i_t - i_0\rangle + \log P_0^{T_i}\}.
\end{aligned} \tag{3.9}$$

(Here we meant $\boldsymbol{\sigma} = (\mu, \sigma)$ and $\boldsymbol{\rho} = (\rho^0, \rho)$.) The last equality, or

$$(T_1 - t)\boldsymbol{\sigma}(T_1, T_2) + \sum_{l=1}^{i-1}(T_{l+1} - T_l)\boldsymbol{\sigma}(T_l, T_{l+1}) = \boldsymbol{\rho}(T_i) - \boldsymbol{\rho}(t), \tag{3.10}$$

is obtained by recalling (3.3).

Therefore we have

$$\begin{aligned}
&(T_l - t - \Delta t)\mathbf{r}_{t+\Delta t}(T_l) - (T_l - t)\mathbf{r}_t(T_l) \\
&\quad = -\langle \boldsymbol{\rho}(T_l) - \boldsymbol{\rho}(t), \Delta i_t\rangle + \langle \boldsymbol{\rho}(t + \Delta t) - \boldsymbol{\rho}(t), i_{t+\Delta t} - i_0\rangle \\
&\quad = -\langle \boldsymbol{\rho}(T_l) - \boldsymbol{\rho}(t), \Delta i_t\rangle + \langle \Delta t\, \boldsymbol{\sigma}(t), i_{t+\Delta t} - i_0\rangle \\
&\quad = -\langle \rho(T_l) - \rho(t), \Delta \mathbf{w}_t\rangle \\
&\qquad\qquad + \big\{\langle \boldsymbol{\sigma}(T_1, T_2), i_{t+\Delta t} - i_0\rangle - \rho^0(T_l) + \rho^0(t)\big\}\Delta t \\
&\quad = -\langle \rho(T_l) - \rho(t), \Delta \mathbf{w}_t\rangle \\
&\qquad + \big\{\langle \sigma(T_1, T_2), \mathbf{w}_{t+\Delta t} - \mathbf{w}_0\rangle - \rho^0(T_l) + \rho^0(t) + \rho^0(t + \Delta t)(t + \Delta t)\big\}\Delta t.
\end{aligned} \tag{3.11}$$

In particular, the left-hand-side of (3.11) has the other of $\sqrt{\Delta t}$. Therefore, the asymptotic expansion with respect to Δt of the present value PV of a cash flow $\mathbf{CF}$ is:

$$
\begin{aligned}
&\frac{\Delta(\mathrm{PV})}{\mathrm{PV}} := \frac{\mathrm{PV}(\mathbf{r}+\Delta\mathbf{r}_t; t+\Delta t) - \mathrm{PV}(\mathbf{r}; t)}{\mathrm{PV}} \\
&= \sum_l \mathbf{CF}(T_l)\left\{e^{-(T_l-t-\Delta t)\mathbf{r}_{t+\Delta t}(T_l)} - e^{-(T_l-t)\mathbf{r}_t(T_l)}\right\}/\mathrm{PV}_t \\
&\simeq -\sum_l \mathbf{CF}(T_l)e^{-(T_l-t)\mathbf{r}_t(T_l)}\left\{(T_l-t-\Delta t)\mathbf{r}_{t+\Delta t}(T_l) - (T_l-t)\mathbf{r}_t(T_l)\right\}/\mathrm{PV}_t \\
&+ \frac{1}{2}\sum_l \mathbf{CF}(T_l)e^{-(T_l-t)\mathbf{r}_t(T_l)}\left\{(T_l-t-\Delta t)\mathbf{r}_{t+\Delta t}(T_l) - (T_l-t)\mathbf{r}_t(T_l)\right\}^2/\mathrm{PV}_t \\
&\qquad + \mathrm{o}(\Delta t).
\end{aligned}
\tag{3.12}
$$

Combining (3.12), (3.11), and (3.10), we have (3.6). To see (3.7) and (3.8) are generalized durations and convexities respectively, just compare (1.9) and (3.6). □

3.3 Another framework based on discrete Itô formula

Since we have assumed *completeness* ($s = n$) in ($a2'$), we can use a discrete Itô formula (c.f.[2, 3]) to obtain perfect equality instead of the above asymptotic expansion (3.6).

Theorem 13. *Define*

$$
\begin{aligned}
\tilde{D}_j := &\frac{e^{\langle\sigma(T_1,T_2),\mathbf{w}_t-\mathbf{w}_0\rangle - \{\rho^0(T_l)-\rho^0(t)\}+\rho^0(t+\Delta t)(t+\Delta t)}}{\Delta t} \\
&\times \mathbf{E}[\Delta w^j e^{\langle\rho(T_l)-\rho(t)+\Delta t\sigma(T_1,T_2),\Delta\mathbf{w}_t\rangle}],\ j = 1, ..., n
\end{aligned}
\tag{3.13}
$$

and

$$
\begin{aligned}
\tilde{D}^2 := \mathbf{E}[&e^{\langle\sigma(T_1,T_2),\mathbf{w}_t-\mathbf{w}_0\rangle - \{\rho^0(T_l)-\rho^0(t)\}+\rho^0(t+\Delta t)(t+\Delta t)} \\
&\times e^{\langle\rho(T_l)-\rho(t)+\Delta t\sigma(T_1,T_2),\Delta\mathbf{w}_t\rangle} - 1]/\Delta t.
\end{aligned}
\tag{3.14}
$$

Then

$$
\frac{\Delta(\mathrm{PV}_t)}{\mathrm{PV}_t} = -\sum_j \tilde{D}_j \Delta w_t^j + \frac{1}{2}\tilde{D}^2\,\Delta t.
\tag{3.15}
$$

Proof. Recall

$$\frac{\Delta(\mathrm{PV}_t)}{\mathrm{PV}_t} = \sum_l \mathbf{CF}(T_l) e^{-(T_l - t)\mathbf{r}_t(T_l)} \left\{ e^{-(T_l - t - \Delta t)\mathbf{r}_{t+\Delta t}(T_l) + (T_l - t)\mathbf{r}_t(T_l)} - 1 \right\} / \mathrm{PV}_t,$$

and

$$\begin{aligned}
&(T_l - t - \Delta t)\mathbf{r}_{t+\Delta t}(T_l) - (T_l - t)\mathbf{r}_t(T_l) \\
&\quad = -\langle \rho(T_l) - \rho(t), \Delta \mathbf{w}_t \rangle \\
&\qquad + \left\{ \langle \sigma(T_1, T_2), \mathbf{w}_{t+\Delta t} - \mathbf{w}_0 \rangle - \rho^0(T_l) + \rho^0(t) + \rho^0(t + \Delta t)(t + \Delta t) \right\} \Delta t.
\end{aligned}$$

We may regard these random variables as a function on S. By putting the latter to be $F(s)$; i.e.

$$\begin{aligned}
F(s) :=& \langle \rho(T_l) - \rho(t) + \Delta t \sigma(T_1, T_2), \Delta \mathbf{w}(s) \rangle \\
&+ \left\{ \langle \sigma(T_1, T_2), \mathbf{w}_t - \mathbf{w}_0 \rangle - \rho^0(T_l) + \rho^0(t) + \rho^0(t + \Delta t)(t + \Delta t) \right\} \Delta t,
\end{aligned}$$

the former is represented as

$$\frac{\Delta(\mathrm{PV}_t)}{\mathrm{PV}_t} = \frac{1}{\mathrm{PV}_t} \sum_l \mathbf{CF}(T_l) e^{-(T_l - t)\mathbf{r}_t(T_l)} \left\{ e^{F(s)} - 1 \right\}. \tag{3.16}$$

Since on the other hand $(1, \Delta w^1/\sqrt{\Delta t}, ..., \Delta w^n/\sqrt{\Delta t})$ is an orthonormal basis of $L(S)$, we have

$$e^{F(s)} = \sum_j \frac{\mathbf{E}[\Delta w^j e^F]}{\Delta t} \Delta w^j + \frac{\mathbf{E}[e^F]}{\Delta t} \Delta t. \tag{3.17}$$

Substituting the right-hand-side of (3.17) for $e^{F(s)}$ in (3.16), we get the expansion (3.15). $\square$

4 Passage to IKRS models

> ... So it was very difficult to estimate the parameters in the original formulation of the mode. That was the motivation for HJM. We looked at the continuous time limit of Ho-Lee... (R. Jarrow [19])

4.1 IKRS interest rate models

In this section we will show that the forward rates of our generalized Ho-Lee model *converge* to those of a special class of HJM which we call IKRS model, as Δt tends to zero.

Let us begin with a quick review of Inui-Kijima-Ritchken-Sankarasubramanian's (IKRS for short) interest rate model [17, 24]. Above all, the class is characterized by the separability of the volatility structure of instantaneous forward rates in a continuous-time HJM framework.

Recall that HJM is based on the semi-martingale representation of the instantaneous forward rate

$$F_t(T) := \lim_{\Delta t \downarrow 0} (\Delta t)^{-1} F_t(T, T + \Delta t).$$

In Brownian cases HJM expression is

$$dF_t(T) = \sum_i v_t^i(T)(dW_t^i - \lambda^i\, dt) + \sum_i \left(v_t^i(T) \int_t^T v_t^i(u)\, du \right) dt. \tag{4.1}$$

Here W^i's are mutually independent Wiener processes and λ^i's are some adapted processes that correspond to so-called *market prices of risk*. IKRS assumes the following separation of volatility:

$$v_t^i(T) := \eta_t^i e^{-\int_t^T \kappa^i(s)\, ds} \quad (i = 1, 2, ..., n), \tag{4.2}$$

where κ^i's are some deterministic functions and η_t^i's are some adapted processes. The favorable property of IKRS lies in the expression of bond prices:

$$P_t^T = \frac{P_0^T}{P_0^t} \exp \sum_i \left\{ -\frac{1}{2} |\Lambda_t^i(T)|^2 \phi_t^i + \Lambda_t^i(T)[F_0(t) - r^i(t)] \right\} \tag{4.3}$$

where

$$\Lambda_t^i(T) = \int_t^T e^{-\int_t^u \kappa_x\, dx}\, du, \tag{4.4}$$

and

$$\begin{aligned} dr_t^i &= [\partial_t F_0(t) + \phi_t^i + \kappa(t)(F_0(t) - r_t^i)]dt + \eta_t^i(dW^i - \lambda^i dt),\ r_0^i = 0 \\ d\phi_t^i &= [|\eta_t^i|^2 - 2\kappa_t \phi_t^i]dt,\ \phi_0^i = 0. \end{aligned} \tag{4.5}$$

The "state variable" $\{r^i, \phi^i\}_{i=1,\dots,n}$ can be, if we impose[5]$\eta^i \equiv \eta(r^i)$, a solution to a degenerate stochastic differential equation under the risk neutral measure, whose density on $\mathcal{F}_t := \sigma(\{W_s^1, ..., W_s^n\}; s \le t)$ with respect to the real world measure is $\mathcal{E}(\sum_i \int \lambda^i dW^i)_t$. For details, see e.g. [5, section 5.3].

Remark 14. The separability of volatility structure is first discussed in [18] by F. Jamshidian, who called the class *Quasi Gaussian*.

4.2 Reparametrizations of IKRS

While the above parameterization of original IKRS is aimed to have Markovian state variables which allow reasonable discretizations, our parameterization below gives a direct discretization scheme by recombining trees though it is limited to the Gaussian cases.

Our new insight starts from the following observation:

Lemma 15. *Let* $\rho^i(t) = \int_0^t e^{-\int_0^s \kappa^i(u)\,du}\,ds$, $\varphi^i(s) = \eta^i(s) e^{\int_0^s \kappa^i(u)\,du}$, *and assume* $\lambda^i(t) = -\rho^i(t)\varphi^i(t)$. *Then the forward rates (4.1) with (4.2) turns into*[6]

$$F_t(T) = F_0(T) + \sum_i \left(\dot{\rho}^i(T) \int_0^t \varphi^i(s)\,dW_s^i + \dot{\rho}^i(T)\rho^i(T) \int_0^t |\varphi^i(s)|^2\,ds \right). \tag{4.6}$$

In particular, when $\varphi^i \equiv 1$, *then* $F_t(T)$ *for each* T *is a process with stationary independent increments.*

Proof. It is direct if one notice that $\dot{\rho}(T) = e^{-\int_0^T \kappa^i(s)\,ds}$ and hence $v_t^i = \eta_t^i e^{-\int_t^T \kappa^i(s)\,ds} = \varphi^i(t)\dot{\rho}(T)$. □

We now consider (4.6) itself to be a new model, meaning that $\dot{\rho}$ can be negative, which is impossible in the original IKRS.

Since the standard forward rate over $[T, T']$ is $\frac{1}{T'-T}(\log P_t^T - \log P_t^{T'}) =: F_t(T, T')$ in the continuous-time framework is represented by the instantaneous ones by

$$F_t(T, T') = \frac{1}{T' - T} \int_T^{T'} F_t(u)\,du.$$

[5] We need to be careful about the linear growth condition in (4.5) when we use such Markovian implementations. For details, see [1].

[6] Here *dot* over a function means its derivative.

Therefore, for a given T, T', IKRS in the form of (4.6) with $\varphi^i \equiv 1$ gives

$$\begin{aligned} F_t(T,T') = F_0(T,T') &+ \left\langle \frac{\rho(T') - \rho(T)}{T' - T}, \mathbf{W}_t - \mathbf{W}_0 \right\rangle \\ &+ \frac{|\rho(T')|^2 - |\rho(T)|^2}{T' - T}\, t. \end{aligned} \tag{4.7}$$

We may extend IKRS to any right continuous ρ by (4.7) if we do not care about the instantaneous forward rates.

4.3 A limit theorem

Comparing (4.7) for $T' = T_{i+1}$ and $T = T_i$ with (3.2)-(3.4), we have the following result.

Theorem 16. *Let t be fixed and $\Delta t = t/N$. The forward rates $\{F_t(T_i, T_{i+1}); i = 1, ..., k\}$ given by (3.2) of the multi-factor generalization of Ho-Lee converge weakly as $N \to \infty$ to the corresponding forward rates of IKRS given by (4.7) for the same ρ.*

Proof. It suffices to show

$$\lim_{\Delta t \to 0} \frac{2}{\Delta t} \log \mathbf{E}[\exp \langle \rho(T), \Delta \mathbf{w} \rangle] = |\rho(T)|^2.$$

But this follows almost directly since we have

$$\begin{aligned} \mathbf{E}[\exp \langle \rho(T), \Delta \mathbf{w} \rangle] &= 1 + \mathbf{E}[\langle \rho(T), \Delta \mathbf{w} \rangle + \frac{1}{2}\mathbf{E}[\langle \rho(T), \Delta \mathbf{w} \rangle^2] + \mathrm{o}(\Delta t) \\ &= 1 + \frac{1}{2}\langle \rho(T) \otimes \rho(T), \mathbf{E}[\Delta \mathbf{w} \otimes \Delta \mathbf{w}] \rangle + \mathrm{o}(\Delta t) \\ &= 1 + \frac{1}{2}\langle \rho(T) \otimes \rho(T), \Delta t I_n \rangle + \mathrm{o}(\Delta t) \\ &= 1 + \frac{1}{2}|\rho(T)|^2 \Delta t + \mathrm{o}(\Delta t). \end{aligned}$$

Here the expectation of a vector is taken coordinate-wisely, and I_n is the $n \times n$ unit matrix. As a matter of course, we have used (1.7). ☐

Remark 17. In this continuous-time framework, the duration-convexity formula like (3.6) can be obtained directly from Itô's formula. In fact, since

$$
\begin{aligned}
-(T_i - t)\mathbf{r}_t(T_i) &= \langle \rho(T_i) - \rho(t), \mathbf{W}_t - \mathbf{W}_0 \rangle + \{|\rho(T_i)|^2 - |\rho(t)|^2\}t + \log P_0^{T_i} \\
&= \sum_j \int_0^t \{\rho^j(T_l) - \rho^j(s)\}\, dW_s^j + \int_0^t \{|\rho(T_i)|^2 - |\rho(s)|^2\}\, ds \\
&\qquad - \sum_j \int_0^t W^j \, d\rho_s - \int_0^t s \, d\mu_s,
\end{aligned}
\tag{4.8}
$$

we have the following equality by applying Itô's formula.

$$
\begin{aligned}
\frac{d(\mathrm{PV})_t}{\mathrm{PV}_t} &= \sum_j \frac{1}{\mathrm{PV}_t} \sum_l (\rho^j(T_l) - \rho^j(t)) \mathbf{CF}(T_l) e^{-(T_l - t)\mathbf{r}_t(T_l)}\, dW^j \\
&+ \frac{1}{2} \sum_{i,j} \frac{1}{\mathrm{PV}_t} \sum_l \{\rho^i(T_l) - \rho^i(t)\}\{\rho^j(T_l) - \rho^j(t)\} \mathbf{CF}(T_l) e^{-(T_l - t)\mathbf{r}_t(T_l)}\, dt \\
&+ \frac{1}{\mathrm{PV}_t} \sum_l \Big(\rho^0(T_l) - \rho^0(t) - t\dot{\rho}^0(t) - \sum_j (W^j - W_0^j)\dot{\rho}^j(t)\Big) \cdot \\
&\qquad \cdot \mathbf{CF}(T_l) e^{-(T_l - t)\mathbf{r}_t(T_l)}\, dt.
\end{aligned}
$$

Thus our generalized durations (3.7) and convexities (3.8) survive the continuous-time framework.

5 Concluding remark

For all refinements cited above, the problem of absurd negative interest rates remains unresolved. To overcome the puzzle, we need to further generalize Ho-Lee model; from linear to non-linear. The results will be published as PART II.

A Appendix

A.1 Principal component analysis: a review

Let $\mathbf{y}_p \equiv (y_p^1, ..., y_p^k)$, $p = 1, 2, ..., N$ be sample data of an $\mathbf{R}^k$-valued random variable $\mathbf{y}$. Define covariance matrix $C := (c_{l,m})_{1 \le l,m \le k}$ by the sample

covariances:

$$c_{l,m} = \frac{1}{N}\sum_{p=1}^{N}(y_p^l - \overline{y}^l)(y_p^m - \overline{y}^m), \tag{A.1}$$

where $\overline{\mathbf{y}}^l = \frac{1}{N}\sum_{p=1}^{N} y_p^l$: the sample average. Since C is a positive definite matrix, it is diagonalized to $\Lambda \equiv \mathrm{diag}[\lambda_1, ..., \lambda_k]$ by an orthogonal matrix $U \equiv [\mathbf{x}_1, ..., \mathbf{x}_k]$; we have $UCU^* = \Lambda$. Here $\mathbf{x}_j \equiv (x_{j,1}, ..., x_{j,k})$ is an eigenvector of the eigenvalue λ_j.

We can assume without loss of generality that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$, and by ignoring small λ's we can choose $\lambda_1, ..., \lambda_n$ (meaning that $\lambda_{n+1}, ..., \lambda_k$ are set to be zero) and corresponding principal component vectors $\mathbf{x}_1, ..., \mathbf{x}_n$. so that we have an expression of

$$\mathbf{y} = \sum_{j=1}^{n} \mathbf{x}_j \sqrt{\lambda_j} w^j + \overline{\mathbf{y}}, \tag{A.2}$$

where $[w^1, ..., w^n] =: \mathbf{w}$ is an $\mathbf{R}^n$-valued random variable with

$$\mathbf{E}[w^j] = 0 \quad \text{and} \quad \mathrm{Cov}(\Delta w^i, \Delta w^j) = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases} \tag{A.3}$$

By redefining $\Lambda^{1/2} := \mathrm{diag}[\lambda_1^{1/2}, ..., \lambda_n^{1/2}]$ as $n \times n$ matrix and $X := [\mathbf{x}_1, ..., \mathbf{x}_n]$ as $k \times n$ matrix, we can rewrite (A.2) as

$$\mathbf{y} = X\Lambda^{1/2}\mathbf{w} + \overline{\mathbf{y}}, \tag{A.4}$$

The expression (A.2)/(A.4) is universal in the following sense: If $\mathbf{y}$ has the expression of

$$\mathbf{y} = A\tilde{\mathbf{w}} + \overline{\mathbf{y}},$$

and if $\tilde{\mathbf{w}} \equiv [\tilde{w}^1, ..., \tilde{w}^n]$ has the first two moments as (A.3), then $k \times n$ matrix A should satisfy $A = X\Lambda^{1/2}T$ for some orthogonal matrix T. In this case $T\tilde{\mathbf{w}}$ again has the moments of (A.3), and so, the expression (A.4) is retrieved.

To be more precise, the above is rephrased as follows; *if the covariance matrix C of $\mathbf{y} - \overline{\mathbf{y}}$ is expressed as AA^*, then $A = X\Lambda^{1/2}T$ for some orthogonal matrix T.* Equivalently, $\Lambda^{-1/2}X^*A$ is an orthogonal matrix. The last statement can be easily checked if one notice that $X\Lambda^{1/2}\Lambda^{1/2}X^* = C$ and that X^*X is the $n \times n$ unit matrix.

A.2 No-arbitrage restriction

The no-arbitrage principle in our framework is, by definition, the impossibility of

$$(-\mathrm{PV}_t, \mathrm{PV}_{t+\Delta t}) > 0 \quad \text{almost surely} \tag{A.5}$$

for any cash flow $\mathbf{CF}$.

Since we have assumed that **there are only finite possibilities** in $(a2)$ or $(a2')$, the inequality can be regarded as the one in a finite dimensional (in fact it is $s+1$ -dimensional) vector space (in the sense that all the coordinates are positive). Therefore, by a separation theorem from the finite-dimensional convex analysis (see e.g. [9, 22]), we know that the impossibility of (A.5) is equivalent to the existence of a strictly positive random variable D_t such that

$$\mathrm{PV}_t(s_{t-1}, ..., s_1) = \sum_{s_t \in S} D_{t+\Delta t}(s_t; s_{t-1}, ..., s_1)\mathrm{PV}_{t+\Delta t}(s; s_{t-1}, ..., s_1), \tag{A.6}$$

for any $(s_{t-1}, ..., s_1)$.

We can reduce the relation (A.6) in the following way.

$$\sum_{T>t} \mathbf{CF}(T) e^{-\mathbf{r}_t(\cdot)(T_l)(T_l - t)} = \sum_{s\in S} D_{t+\Delta t}(s; \cdot) \sum_{T>t} \mathbf{CF}(T) e^{-\mathbf{r}_{t+\Delta t}(T)(s;\cdot)(T_l-(t+\Delta t))}$$

$$\text{for any flow } \mathbf{CF},$$

$$\begin{aligned} &\Longleftrightarrow\ e^{-\mathbf{r}_t(\cdot)(T)(T-t)} = \sum_{s\in S} D_{t+\Delta t}(s;\cdot) e^{-\mathbf{r}_{t+\Delta t}(T)(s;\cdot)(T-(t+\Delta t))} \\ &\Longleftrightarrow\ P_t^T(\cdot) = \sum_{s\in S} D_{t+\Delta t}(s;\cdot) P_{t+\Delta t}^T(s;\cdot) \end{aligned} \tag{A.7}$$

By introducing

$$\delta_t(s;\cdot) := D_t(s;\cdot) / \Pr(\Delta \mathbf{w}_t = s | \mathcal{F}_{t-\Delta t})(\cdot),$$

we can rewrite the last equation in (A.7) as

$$P_t^T = \mathbf{E}[\delta_{t+\Delta t} P_{t+\Delta t}^T | \mathcal{F}_t], \tag{A.8}$$

where $\mathcal{F}_t = \sigma(s_t, s_{t-1}, ..., s_1)$. By the tower property of conditional expectation we have

$$P_t^T = \mathbf{E}[\prod_{u=t+\Delta t}^{T} \delta_u | \mathcal{F}_t], \tag{A.9}$$

and in particular

$$P_t^{t+\Delta t} = \mathbf{E}[\delta_{t+\Delta t}|\mathcal{F}_t].$$

By introducing

$$\begin{aligned}\pi_t(s;\cdot) &:= \frac{\delta_t}{\mathbf{E}[\delta_{t+\Delta t}|\mathcal{F}_t]}\Pr(\Delta\mathbf{w}_{t+\Delta t} = s|\mathcal{F}_t)(\cdot) \\ &= \frac{\delta_t}{P_t^{t+\Delta t}}\Pr(\Delta\mathbf{w}_{t+\Delta t} = s|\mathcal{F}_t)(\cdot) = D_t(s;\cdot)/P_t^{t+\Delta t},\end{aligned}$$

we have

$$P_t^T = \sum_{s\in S}\pi_{t+\Delta t}(s)P_t^{t+\Delta t}P_{t+\Delta t}^T. \tag{A.10}$$

This means that π is a risk neutral probability in the standard terminology: Defining $H_{t+\Delta t}(T) := P^{(t+\Delta t)}(T)P^{(t)}(1)/P^{(t)}(T+1)$, (A.10) is equivalent to

$$\mathbf{E}^\pi[H_{t+\Delta t}|\mathcal{F}_t] = 1, \tag{A.11}$$

where $\mathbf{E}^\pi$ stands for expectation with respect to π.

It is easy to see that the converses, starting from existence of π, ending in that of D, are true.

To summarize, we have the following.

Theorem 18. *The following statements are equivalent.*

(i) The market is arbitrage-free.

(ii) There exists a strictly positive adapted process $\{D_t\}$ such that (A.7) holds.

(iii) There exists a strictly positive adapted process $\{\delta_t\}$ such that (A.9) holds.

(iv) There exists a risk neutral probability π (such that (A.10) or (A.11) holds).

Now we give a proof of Theorem 8.

Proof of Theorem 8. Suppose we have a state price density D. Define $\delta_t := D_t/D_{t-1}$. Then (A.9) holds. The converse is trivial. □

References

[1] Akahori, J.: Explosion tests for stochastic integral equations related to interest rate models, *J. Math. Sci. Univ. Tokyo* **5** (1998) 727–745.

[2] Akahori, J.: A discrete Ito calculus approach to He's framework for multi-factor discrete market, preprint, Ritsumeikan University , 2004.

[3] Akahori, J.: Discrete Ito Formulas and Their Applications to Stochastic Numerics, *RIMS kokyuroku* **1462** 202-210 (2006) ; proceedings of the 7th Workshop on Stochastic Numerics; Jun 27–29, 2005, RIMS, Kyoto. arXiv math.PR/0603341

[4] Brace, A., Gatarek, D. and Musiela, M.: The market model of interest rate dynamics, *Mathematical Finance* **7** (2) (1997) 127-155.

[5] Brigo, D., and Mercurio, F.: *Interest Rate Models—Theory and Practice.* Springer Finance. Springer-Verlag, Berlin, 2001

[6] Campbell, J.Y., Lo, A.W. and MacKinlay, A.C.: *The Econometrics of Financial Markets*, Princeton University Press, 1997.

[7] Chen, L., D. Filipović, and H. V. Poor: Quadratic term structure models for risk-free and defaultable rates, **Math. Finance**, 14(4), (2004) 515-536.

[8] Cox, J.C., J.E. Ingersoll, and S.A. Ross: A Theory of the Term Structure of Interest Rates, *Econometrica* **53** (1985) 385-408.

[9] Duffie, D.: *Dynamic Asset Pricing Theory*, 3rd eds. Princeton University Press, 2001.

[10] Duffie, D. and Kan, R.: A Yield Factor Model of Interest Rates, *Mathematical Finance* **6**(4), (1996) 379-406.

[11] Heath, D., R.A. Jarrow and A. Morton: Bond Pricing and the Term Structure of Interest Rates: A Discrete Time Approximation. *Journal of Financial Quantitative Analysis* **25** (1990), 419-440.

[12] Heath, D., R.A. Jarrow and A. Morton: Bond Pricing and the Term Structure of Interest Rates: A New Methodology. *Econometrica* **60** (1992), 77-105.

[13] Ho, T.S.Y., and S. B. Lee: Term-structure movements and pricing interest rate contingent claims, *Journal of Finance* **41** (1986) 1011–1029.

[14] Ho, T.S.Y., and S. B. Lee: A Closed-Form Multifactor Binomial Interest Rate Model Ho, Thomas S.Y. Lee, Sang Bin *The Journal of Fixed Income* **14-1**, (2004) 8–16.

[15] Hull, J.C.: *Options, Futures, and Other Derivatives*. 6th ed. Prentice-Hall International Editions. Upper Saddle River, NJ: Prentice Hall. 2003.

[16] Hull, J. and A. White: Pricing Interest-Rate Derivative Securities, *Review of Financial Studies*, **3** (1990), 573–592.

[17] Inui, K. and Kijima, M.: A Markovian framework in multi-factor Heath-Jarrow-Morton models, *Journal of Financial and Quantitative Analysis*, **33** (3), (1998), 423 -440.

[18] Jamshidian, F.: Bond and option evaluation in the Gaussian interest rate model, *Research in Finance*, **9** (1991), 131–170.

[19] Jarrow, R. interview in *Derivatives Strategy*, April, 1997. (available at `http://www.derivativesstrategy.com/magazine/archive/1997/0497qa.asp`)

[20] Littermanm, R. and J. Scheinkman: Common factors affecting bond returns *The Journal of Fixed Income* **1** (1988), 54–61.

[21] Longstaff, F and E. S. Schartz: Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model, *The Journal of Finance* **47** (1992) 1259-1282.

[22] Rockefeller, RT: *Convex Analysis*. Princeton University Press, 1970.

[23] Ross, S.A.: The arbitrage theory of capital asset pricing, *Journal of Economic Theory*, **13** (1976) 341-360.

[24] Ritchken P. and L. Sankarasubramanian: Volatility structure of forward rates and the dynamics of the term structure, *Mathematical Finance* **5** (1995) 55–72.

[25] Vasicek, O. : An equilibrium characterisation of the term structure. *Journal of Financial Economics* **5** (1977) 177–188.